Multisubmodular Function Based Partitioning for Tradeoff between Distinguishability, Binning Cost and Meta Information Protection

Farhang Bayat, Shuangqing Wei
Department of Electrical and Computer Engineering, Louisiana State University, USA

Abstract—**In this paper, we propose a novel formulation to understand the tradeoff between binning cost, meta information leakage and hypotheses distinguishability in communication networks. To do so, we reiterate the idea of partitioning random items from our previous work.**

Under this framework, the goal is to maximize the measure of distinguishability between two hypotheses while restraining a certain level of information leakage and overall binning cost by properly dividing a set of M random items into N bins.

To do so, we formulate a novel multi-agent multi-variate optimization problem which are generally NP-complicated. We then utilize the submodular nature of the problem to find sufficient conditions to (1) secure the existence of a solution to our problems; (2) lower the complexity of the problem at the cost of some accuracy. After offering sufficient conditions to secure the existence of a solution in all general cases, we offer examples which help signify the importance of using submodular properties.

Index Terms—**partition, information leakage, privacy, mutual information, hypotheses distinguishability, submodularity**

I. Introduction

The concept of a tradeoff between an overall binning function and meta information leakage in a communication network is one of the most recurring subjects in modern engineering research. In this paper, we aim to offer a novel insight into and a novel formulation of such problems. To motivate such concerns, we choose to first provide an example about the necessity of considering such tradeoff.

Through our use of internet, there are many times when we wish to be distinguished from other users (so as to enjoy a tailor-made service) while limiting the amount of information reflected to an eavesdropper by patterns shown in sequence of webpages we have visited.

One method to carry this out, is the use of Virtual Private Network (VPN)s. Such networks offer the user complete anonymity by allocating a fake IP address to the user rendering him untraceable to the eavesdropper. Thus, if we could address all our browses through the same VPN, no information would be leaked. However, a VPN tends to slow down the connection speed . Also, the pace is further dropped when more traffic (other users' data) is imposed upon the network. As a solution, we can employ multiple VPNs and thus cut down on the utility loss. However, by using multiple VPNs we are allowing a level of privacy breach into our browses where the eavesdropper is able to deduce some information about our tendencies by observing the distribution of used webpages over multiple networks. Then, if a utility function based upon connection speed -bandwidth- for the user is calculable, a privacy constrained problem between the user and an eavesdropper (for example a service provider) could be defined.

The solution to such a problem could offer insights in regards to the tradeoff between VPN allocation utility and meta information leakage when we face the problems of partitioning a set of random items (i.e. websites that a user has chosen to visit following his own distributions) into a given number of bins (i.e. a given set of VPN servers each of which has its own utility function, as will be further detailed in Section II).

In our proposed framework as detailed in Section II, meta data information refers to the patterns about a sequence of items (for example the user's favorable websites) infer-able based upon a sequence of bins (e.g. proxy sites) observed by an eavesdropper. This assumption is an expansion of [1] and [2] and our own work in [3].

Furthermore, in our investigation we find it necessary to introduce a measure of distinguishability and a binning cost function.

Finally, we introduce multi-submodularity and submodularity as two means of reducing the complexity level of such problems, namely, dividing M random items into N bins, under an upper-bound on leaked meta information.

The concept of privacy has already been explored in many works such as [4], [5] where a general but non-mathematical explanation was offered. However; in our work, we go into further details as to what privacy represents in our framework and how it could be formulated into many settings. Later, we find it necessary to utilize the concept of multi-submodular set function problems and their solutions. This concept was widely discussed in [6] where they introduced a series of sufficient conditions on multi-submodular set functions by which the multi-submodular problem could be transformed into a submodular set function problem. Then, further discussions about the existence of a solution to the new problem were made. By doing so -and if a solution were proven to exist-, the complexity of the problem could be shown to be reduced from NP to polynomial. However, [6] did not offer any algorithmic solutions in such cases which unfortunately results in us simply proving the existence of a solution rather offering an algorithm to support such solution as well.

As for the addition of distinguishability, to the best of our knowledge, the closest work to ours was done by [7] where

they considered the tradeoff between distinguishability and information leakage when the former is quantified using KL-distance and the latter using mutual information.

It is important to note that some of the concepts revisited in this paper, were already well-defined in our previous work [3]. Thus, in this paper, we refer the reviewer to [3] for some problem formulations and instead attempt to address the problem concerning distinguishability in full details.

As for the novel problem addressed in this paper, our goal is to offer (1) a measure of distinguishability between the K hypothesis; (2) a measure of average leaked information given any of the K hypothesis is active; (3) a multi-agent multi-variant optimization problem with privacy leakage and overall binning cost constraints; (4) insight into the complexity of such an NP-hard problem and assumptions we may need to incur to simplify it into a polynomial problem such as aiming to minimize the upper bound or the form of the cost function; (5) a full description of the algorithm utilized to find the solution given the sufficient conditions followed by the proximity results and finally (6) examples to further demonstrate the applicability of submodular solutions in our own and even more general problems of such nature.

The rest of this paper is organized as follows. In Section II we formulate the two problems in terms of privacy and utility functions. We dissect what the goal and the constraints are. In Section III, we inspect the overall utility functions of each problem and then find the sufficient conditions under which they are equipped with multi-submodular property. Due to general limited knowledge about multi-submodular solutions and how they are developed, we then assume a specific case $N = 2$. In Section IV two algorithms for each case are developed which offer a polynomial complexity and an accuracy level of e or $\frac{1}{e}$ depending on whether the goal is to minimize or maximize the overall utility function respectively.

II. System Model

In this section, we aim to formulate the problem of distinguishing two possible hypotheses while keeping the average leaked information and the overall binning cost to two specific thresholds. To do so, we assume access to two multinomial distributions G_1 and G_2 each with prior probabilities $P(G_1)$ and $P(G_2) = 1 - P(G_1)$ defined over M items. These multinomial distributions impose probabilities π_{i_p} over item i given multinomial distribution G_p is chosen where $p \in \{1, 2\}$, $i \in \{1, \cdots, M\}$. It also follows that $\sum_{i=1}^{M} \pi_{i_p} = 1, p = 1, 2$.

We then aim to distinguish these two distributions by choosing binning as our method of observation meaning we attempt to find an M-to-N mapping between the set of M items and set of N outputs which can best help us decide which multinomial distribution was chosen.

We further assume that binning the M items is a costy function whose amount of cost is based upon the sum of overall probabilistic weight of items mapped to each output respectively. We further note that by imposing a binning procedure over the set of M items we are leaking a certain amount of information about them (seeing as how frequencies of their activity is exposed).

A. Set Allocation

First, we propose an abstract framework to formalize the goal of seeking partition of M items into N bins. More specifically, we aim to allocate each one of $1 \leq i \leq M$ possible items to one of N output bins. There could be at most N^M such partitions. It follows that any set allocation $A_l, 1 \leq l \leq N^M$ results in N sets $S_j^{(l)} \subseteq \{1, 2, ..., M\}, j = 1, 2, ..., N$. Each such set is defined as

$$S_j^{(l)} = \{i | \theta_{i,j}^{(l)} = 1\}$$
$$where \quad \theta_{ij}^{(l)} = \begin{cases} 1 & i \in S_j^{(l)} \\ 0 & i \notin S_j^{(l)} \end{cases} \tag{1}$$

We further assume $S_j^{(l)} \cap S_k^{(l)} = \emptyset, j \neq k$. Furthermore we have $\bigcup_{j=1}^{N} S_j^{(l)} = \{1, 2, ..., M\}$. Finally the size of each such set $S_j^{(l)}$ is defined as $L_j^{(l)}$.

B. Probabilistic Model

We assume at any time slot one and only one of the inputs is chosen with a certain probability. Thus, assuming hypotheses G_p is active and if we use variable $X \in \{1, 2, ..., M\}$ as a representation of set of items, we could have $P(X = i | G_p) = P(\gamma_{i_p} = 1) = \pi_{i,p}, 1 \leq i \leq M, , p \in \{1, 2\}$ as a representation of the probability of choosing item i from the set X under hypotheses G_p where $\gamma_{i_p} \in \{0, 1\}$. It further follows that $\sum_{i=1}^{M} \gamma_{i_p} = 1, p \in \{1, 2\}$, stipulating that one and only one of M items is selected.

Next, we introduce an observable random variable $Y \in \{1, 2, \cdots, N\}$, denoting the index of the bin (the proxy site) employed to carry one of the $M > N$ items. It follows that the probability of each bin's appearance given a set allocation scheme such as A_l under hypotheses G_p will be equal to

$$\begin{aligned} &P(Y = j | A_l, G_p) \\ &= \sum_{i=1}^{M} P(Y = j | A_l, X = i, G_p) P(X = i | A_l, G_p) \\ &= \sum_{i=1}^{M} P(Y = j | A_l, X = i) P(X = i | G_p) \end{aligned} \tag{2}$$

Furthermore, $P(Y = j | A_l, X = i) = \theta_{ij}^{(l)} \in \{0, , 1\}$. It thus follows that

$$P(Y = j | A_l, G_p) = \sum_{i=1}^{M} \theta_{ij}^{(l)} P(X = i | G_p) \rightarrow$$
$$P(Y = j | A_l, G_p) = \sum_{i \in S_j^{(l)}} \pi_{i_p} = \alpha_j^{(l,p)} \tag{3}$$

C. Revealed Information

By choosing to allocate M items to N bins where $N \leq M$, we have injected ambiguity and uncertainty into the output binning index sequence about the input item sequence over a successive n visits or channel uses. In other words, if we originally chose to transmit n of such items, our total set of possible sequences would be of form $\overrightarrow{\mathbf{X^n}} = [X_1 X_2 ... X_n]$ out of M^n possible outcomes. From an observer's perspective which can only have access to which one of N bins is deployed in each time slot, sequences in the form of $\overrightarrow{\mathbf{Y^n}} = [Y_1 Y_2 ... Y_n]$ has cardinality of at most $N^n < M^n$. Despite the amount of uncertainty added due to the many-to-one binning, the output sequence sill reveals certain amount of information regarding the patterns of sequences of M random items.

This observation could be further studied by indicating how our allocation system resembles a coding framework where we have an equivalent channel whose input variable is X and output Y, as demonstrated in Figure 1.

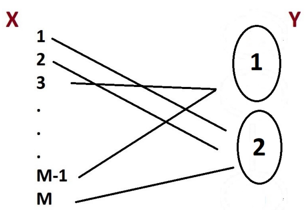

Figure 1. Coding Channel Representation of the Problem

Under such a framework, the equivalent channel output sequence $\overline{\mathbf{Y}}^n$ can help an eavesdropper classify the input sequence $\overline{\mathbf{X}}^n$ into a number of differential classes. As a result, information about the specific input item patterns is leaked to certain degree and can be measured using conditional mutual information $I(X;Y|A_l, G_p)$ between X and Y, under a hypotheses G_p given a particular binning (i.e. partition A_l relationship as illustrated in Figure 1.

Such conditional mutual information thus measures the maximum number of bits of meta information about item sequence per channel use. Therefore, we can have at most $2^{nI(X;Y|A_l,G_p)}$ sequences $\overline{\mathbf{X}^n}$ distinguishable by inferring based on $\overline{\mathbf{Y}^n}$. We thus adopt $I(X;Y|A_l, G_p)$ as the privacy metric conditioned on a particular partition mapping A_l under a specific hypotheses G_p. It follows that due to the combinatorial nature of a set allocation problem there are a total of N^M possible methods to allocate these M items to the N sets.

We can formulate the mutual information over a set allocation $A_l, 1 \leq l \leq N^M$ under hypotheses G_p as:

$$I(X;Y|A_l,G_p) = H(X|A_l,G_p) - H(X|Y,A_l,G_p) = H(Y|A_l,G_p) - H(Y|X,A_l,G_p) = H(Y|A_l,G_p) = H(\alpha_1^{(l,p)}, \alpha_2^{(l,p)}, ..., \alpha_N^{(l,p)}) \quad (4)$$

where we have used the notion of $H(a_1, a_2, ..., a_m) = -\sum_{v=1}^{m} a_v \log a_v$ and the fact that $H(Y|X, A_l, G_p) = 0$ because if the input X and the channel scheme A_l and hypotheses G_p are known, then output Y will offer no uncertainty.

D. Distinguishability Measure

In this section, we aim to define a measure which can evaluate the distinguishability between a number of multinomial distributions on M items.

Given that a certain set allocation A_l has been chosen, depending on the choice of G_p, two different distributions defined by using variables $\alpha_j^{(l,p)}$ could be developed where we have:

$$\alpha_j^{(l,p)} = P(Y = j|A_l, G_p), p \in \{1,2\}, j \in \{1, ..., N\} \quad (5)$$

where the two probability sets $P_1 = \{\alpha_j^{(l,1)}, j = 1, 2, ..., N\}$ and $P_2 = \{\alpha_j^{(l,2)}, j = 1, 2, ..., N\}$ represent the multinomial distributions under allocation scheme l given two prior probability distributions G_1 and G_2 which we aim to distinguish. To evaluate the distinguishability between two such distributions, we utilize the definition of symmetric KL divergence:

$$\frac{1}{2}\{KL(P_1||P_2) + KL(P_2||P_1)\} = \frac{1}{2}\sum_{j=1}^{N}(\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) \quad (6)$$

E. Cost Function

In the problem of hypotheses distinguishability, we note that we are utilizing binning as a means of detecting the multinomial distribution imposed upon the items. It thus makes sense that binning would be a costy process. We assume that the cost function is defined as the probabilistic sum of respective set functions based upon items allocated to each bin. In other words, we assume given a set allocation A_l and that G_p is selected, the cost function follows the following format

$$C(A_l, G_p) = \sum_{j=1}^{N} \alpha_j^{(l,p)} f(S_j^{(l,p)}) \quad (7)$$

where $S_j^{(l,p)}$ represents the set of items allocated to bin j given that hypotheses G_p is active. We thus aim to minimize the average of $C(A_l, G_p)$.

$$C(A_l) = \sum_{p=1}^{2} P(G_p) \sum_{j=1}^{N} \alpha_j^{(l,p)} f(S_j^{(l,p)}) \quad (8)$$

F. Problem Formulation

As was mentioned in the beginning of this section, our overall goal could be summed up in finding a partition of M items into N bins under which we can maximize the distinguishability between hypotheses G_1 and G_2 while keeping the overall leaked information and cost function to a threshold. We thus formulate the following constrained optimization problem:

$$\min_{1\le l\le N^M} \quad \sum_{p=1}^{2}\sum_{j=1}^{N} P(G_p)[\alpha_j^{(l,p)} f(S_j^{(l,p)}) -\lambda_2 \alpha_j^{(l,p)} \log \alpha_j^{(l,p)}] -\frac{\lambda_1}{2}\sum_{j=1}^{N} (\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) \quad (9)$$

while making certain that

$$\sum_{j=1}^{N}\sum_{p=1}^{2} P(G_p)[-\alpha_j^{(l,p)} \log(\alpha_j^{(l,p)})] \le I_{th} \quad (10)$$

and

$$\sum_{p=1}^{2}\sum_{j=1}^{N} P(G_p)\alpha_j^{(l,p)} f(S_j^{(l,p)}) \le \mathcal{C} \quad (11)$$

hold true. It is important to note that the variables λ_1 and λ_2 represent the weight offered to each constraint in comparison with the objective function. Thus, the problem formulated in Eq.(9) is not a Lagrange multiplier problem where we aim to find the optimal values for λ_1 and λ_2; they are simply given to us by the user.

III. Multi-Submodular Set Functions as a Means of Solution

Unfortunately, the problem formulated in Eq.(9) is NP-complicated (it is only solved when a search over N^M possible set allocations is fully carried out and the optimal set allocation is revealed). Still, we could opt to utilize the definition of multi-submodular set functions so as to reduce the complexity to that of polynomial at the cost of accuracy.

In other words, we aim to introduce a model where the user inputs prior probability distribution of different hypotheses, item distribution under each hypotheses, the cost set function, the leaked information threshold I_{th}, the overall binning cost threshold $\mathcal{C}$ and the constraint weights λ_1 and λ_2 and we decide whether a semi-optimal set allocation scheme for such a setting could be found.

A. Imposing Multi-submodularity

In [6], it was shown that if we can prove multi-submodularity for functions such as those formulated in Eq. (9), then they could be modeled as simpler problems (submodular set functions). We thus, aim to find the sufficient conditions for such occurrence. To do so, we first offer a review of multi-submodularity.

As mentioned in [6], if we define $\mathbb{M} = \{1, 2, ..., M\}$, then a multivariate function $F : (2^{\mathbb{M}})^N \to \mathbb{R}^+$ is multi-submodular if for all pairs of tuples $(S_1, ..., S_N)$ and $(T_1, ...T_N) \in (2^{\mathbb{M}})^N$ we will have:

$$F(S_1,...,S_N) + F(T_1,...,T_N) \ge F(S_1\cup T_1,...,S_N\cup T_N) + F(S_1\cap T_1,...,S_N\cap T_N) \quad (12)$$

Since in our formulation functions are separately defined on different sets, the condition in Eq. (12) is simplified to the sufficient condition of submodularity of:

$$F(S_j^{(l)}) = P(G_1)[\alpha_j^{(l,1)} f(S_j^{(l,1)}) - \lambda_2\alpha_j^{(l,1)} \log \alpha_j^{(l,1)}] + P(G_2)[\alpha_j^{(l,2)} f(S_j^{(l,2)}) - \lambda_2\alpha_j^{(l,2)} \log \alpha_j^{(l,2)}] - \frac{\lambda_1}{2}(\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) = F(S_j) \quad (13)$$

for all sets S_j. Finally, in order to further simplify the formulation of the problem we aim to drop l from the above definition and rewrite:

$$F(S_j) = P(G_1)[\alpha_j^{(1)} f(S_j^{(1)}) - \lambda_2\alpha_j^{(1)} \log \alpha_j^{(1)}] + P(G_2)[\alpha_j^{(2)} f(S_j^{(2)}) - \lambda_2\alpha_j^{(2)} \log \alpha_j^{(2)}] - \frac{\lambda_1}{2}(\alpha_j^{(1)} - \alpha_j^{(2)})(\log \alpha_j^{(1)} - \log \alpha_j^{(2)}) \quad (14)$$

The new denotation alludes to the fact that once a set allocation A_l is chosen, its index could be dropped.

In the next step, we opt to use diminishing return property as the means of making certain each of these functions is submodular. Following is a definition of diminishing returns for submodular functions.

Diminishing Property Return dictates that if we define $\mathbb{S}$ as the universal set, a set function $F : 2^{\mathbb{S}} \to \mathbb{R}^+$ is submodular if, for all $A, B \subseteq \mathbb{S}$ with $A \subseteq B$ and for each $x \in \mathbb{S} - B$ we have [8]:

$$F(A \cup \{x\}) - F(A) \ge F(B \cup \{x\}) - F(B) \quad (15)$$

Note: For any further references, we first need to address a series of variable and function definitions which are going to play a vital role in the rest of this chapter:

Definitions for Problem 1

1. Any variable represented with a capital Letter represents a set.
2. Any variable represented with a small letter represents an element.
3. $A - B$ represents a set containing all elements of set A which do not appear in set B.
4. α_x represents the probability of item x and α_A represents the sum of probabilities of items mapped into a set A.
5. $\alpha_{B\setminus A}$ represents the difference in the sum of probabilities of items mapped into the sets B and A which could be further shown as $\alpha_{B\setminus A} = \alpha_B - \alpha_A$.
6. $g(C, D)$ represents the 1^{st} order difference of a set function $f(C)$ from $f(C - D)$ where $D \subseteq C$ which could be formulated as $g(C, D) = f(C) - f(C - D)$.
7. $q(C, C_1, D, D_1)$ represents the 2^{nd} order difference of a set function $f(C)$ where $C_1 \subseteq C$ and $D_1 \subseteq D$ which could be formulated as $q(C, C_1, D, D_1) = g(C, D) - g(C_1, D_1)$.
8. We assume the probability of items is sorted in a decreasing manner such as $\pi_{1_1} \ge \pi_{2_1} \ge ... \ge \pi_{M_1}$ over Hypotheses G_1 and $\pi_{1_2} \ge \pi_{2_2} \ge ... \ge \pi_{M_2}$ over Hypotheses G_2.

Next, we find sufficient conditions for multisubmodularity of the overall utility function described in Eq. (9).

At a first glance, it seems that the function $F(S_j) = F(S_j^{(l)})$ has a singular relationship with set $S_j^{(l)}$. However;

there is a secondary relationship the function shares with the set $S_j^{(l)\mathbf{C}} = \mathbb{S} - S_j^{(l)}$ where $\mathbb{S}$ represents the universal set. This relationship could be modeled as

$$\begin{aligned} F(S_j^{(l)\mathbf{C}}) &= P(G_1)[(1-\beta_j^{(l,1)})f(S - S_j^{(l,1)\mathbf{C}}) \\ &-\lambda_2(1-\beta_j^{(l,1)})\log(1-\beta_j^{(l,1)})] \\ &+P(G_2)[(1-\beta_j^{(l,2)})f(S - S_j^{(l,2)\mathbf{C}}) \\ &-\lambda_2(1-\beta_j^{(l,2)})\log(1-\beta_j^{(l,2)})] \\ -\frac{\lambda_1}{2}&(\beta_j^{(l,2)} - \beta_j^{(l,1)})(\log(1-\beta_j^{(l,1)}) - \log(1-\beta_j^{(l,2)})) \end{aligned} \tag{16}$$

where we have used the fact that $\beta_j^{(l)} = 1 - \alpha_j^{(l)}$ seeing as how we define

$$\beta_j^{(l)} = \sum_{i \in \{S - S_j^{(l)}\}} \pi_i \tag{17}$$

As was the case for Eq.(14), we choose to simplify Eq.(16) in the following manner:

$$\begin{aligned} F(S_j^{\mathbf{C}}) &= P(G_1)[(1-\beta_j^{(1)})f(S - S_j^{(1)\mathbf{C}}) \\ &-\lambda_2(1-\beta_j^{(1)})\log(1-\beta_j^{(1)})] \\ &+P(G_2)[(1-\beta_j^{(2)})f(S - S_j^{(2)\mathbf{C}}) \\ &-\lambda_2(1-\beta_j^{(2)})\log(1-\beta_j^{(2)})] \\ -\frac{\lambda_1}{2}&(\beta_j^{(2)} - \beta_j^{(1)})(\log(1-\beta_j^{(1)}) - \log(1-\beta_j^{(2)})) \end{aligned} \tag{18}$$

Thus, for any set S_j, we must find the sufficient conditions for the existence of diminishing property for both functions described in Eq.(14) and Eq.(18). To do so, we will evaluate their sufficient conditions and then find their intersection as the final conditions (assuming they do not negate one another).

Theorem III.1. *The objective function defined in Eq.(9) is submodular if*

$$\begin{aligned} &(1) \quad g(S_j, S_w) \le 0 \\ &(2) \quad q(S_j, S_w, S_r, S_t) \le 0 \\ &(3) \quad |q(S_j, S_w, S_r, S_t)| \ge max(Z_1, Z_2, Z_3, Z_4), \\ &Z_1 = \lambda_2 \log \omega' + \frac{\lambda_1}{P(G_1)} \log(\omega'\chi'), \\ &Z_2 = \lambda_2 \log \chi' + \frac{\lambda_1}{P(G_2)} \log(\omega'\chi'), \\ &Z_3 = \lambda_2 \log \omega_2' + \frac{\lambda_1}{P(G_1)} \log(\omega_2'\chi_2'), \\ &Z_4 = \lambda_2 \log \chi_2' + \frac{\lambda_1}{P(G_2)} \log(\omega_2'\chi_2'), \\ &\omega' = 1 + \frac{\pi_{1_1}}{\pi_{M_1}(\pi_{1_1} + \pi_{M_1})}, \\ &\chi' = 1 + \frac{\pi_{1_2}}{\pi_{M_2}(\pi_{1_2} + \pi_{M_2})}, \\ &\omega_2' = 1 + \frac{\pi_{1_1}}{1 - \pi_{1_1}}, \\ &\chi_2' = 1 + \frac{\pi_{1_2}}{1 - \pi_{1_2}} \end{aligned} \tag{19}$$

for all possible sets $S_w \subseteq S_j$ and $S_t \subseteq S_r$ and $S_r \subseteq S_j$.

The proof for this theorem is presented in the appendix under Theorem III.1. The proof simply consists of writing down the diminishing returns property inequality and finding sufficient conditions so that each positive value's coefficient is also positive.

IV. Submodular Solution

In this section, we offer an algorithm specifying how the problem formulated in Eq.(9) could be solved in a polynomial manner. It is important to note that this algorithm only offers a solution for when $N = 2$. However, even in the case of a binary output, an exhaustive search method would result in a 2^M complex problem. We thus, first offer the algorithm and then expand upon how the complexity of the solution has been compromised. Finally, we offer an example to showcase the relationship between solution error and complexity.

A. Solution to Eq.(9)

To offer our algorithm we define the objective argument in Eq.(9) as $T(S_j^{(l)})$ and the constraint inequalities in Eq.(10) and Eq.(11) as $U(S_j^{(l)})$ and $W(S_j^{(l)})$.

We then can write:

Algorithm 1: Submodular Function Solution to the problem as described in Eq.(9)

1. Let $S_1 = argmax_{e \in X = \{1,\ldots,M\}} T[S_1 = \{e\}]$ while $U(S_1 = \{e\})$ and $W(S_1 = \{e\})$ both satisfy the constraints.
2. If there is an element $e \in X \setminus S_1$ such that $T[S_1 + \{e\}] \ge T[S_1]$ and $U(S_1 + \{e\})$ and $W(S_1 + \{e\})$ both satisfy the constraints, let $S_1 = S_1 + \{e_1\}$.
3. If there is an element $e \in S_1$ such that $T[S_1 \setminus \{e\}] \ge T[S_1]$ and $U(S_1 - \{e\})$ and $W(S_1 - \{e\})$ both satisfy the constraints, let $S_1 = S_1 - \{e\}$. Go to Step 2.
4. Return maximum of $T[S_1]$ and $T[X \setminus S_1]$.

Here, we know that at the very last step $T[S_1] = T[X \setminus S_1]$. Now we opt to calculate the complexities of this method. Steps 2 and 3 could repeat $(M-1) + (M-2) + ... + 1 = \frac{M(M+1)}{2}$ times each while every item could be removed and thus replaced a total of $2M$ times. Thus the total complexity of steps 2 and 3 is equal to $M^2(M+1) = O(M^3)$. The complexity of step 1 is also equal to M. Thus the total complexity of the solution is equal to $O(M^3)$.

This polynomial solution simply makes certain the maximal and minimal functions obtained are at least 0.432 and at most 2.315 times the optimal objective functions respectively. This range of error occurs because in this method, we are removing and adding members from and to the set S_1 one by one. Thus, at each decision point we are making one locally optimal decision. However, it is widely known that a locally greedy method is not necessarily globally optimal [9].

V. Numerical Examples

We aim to offer the reader two problems where we are hoping to use the results gathered in Theorem III.1 to first find a proper utility function given the specifics of each case. We will then follow Algorithm 2 and compare

its results with that of an exhaustive search to compare the two methods in terms of complexity and exactness of the solution in both problems. In both examples, we assume $M = 6, N = 2, \lambda_1 = 10, \lambda_2 = 0.1, I_{th} = 0.8, \mathcal{C} = 3000$ and that $f(S) = f(|S|)$ where $|S|$ represents the cardinality of a set S. Furthermore we assume that $P(G_1) = 0.4, P(G_2) = 0.6$. The only difference between the two examples would then lay in their respective prior item distributions.

Note: In our following derivation of a desirable cost function we focus on developing a quadratic cost function. Such an assumption might appear to be extremely limiting to our class of functions at first sight. However, as is further explained in [10], such an assumption is quite understandable and rather desirable seeing as how in many economic models, functions are written as extensions of quadratic function and thus later calculations are simplified while still maintaining the essence of a cost function.

A. *Cost function for Table I*

Table I
EXAMPLE 1 FOR ITEM PROBABILITY DISTRIBUTION

Hypotheses	e_1	e_2	e_3	e_4	e_5	e_6
H_1	0.11	0.11	0.09	0.10	0.10	0.49
H_2	0.01	0.01	0.51	0.02	0.03	0.42

One of the simplest utility functions which could satisfy the conditions presented in Theorem III.1 is a quadratic function in the form of $f(S) = a|S|^2 + b|S| + C$. By calculating the 1^{st} and 2^{nd} order differences we can see that they follow the form of $g(|S|) = 2a|S| + b - a$ and $q(S) = 2a$ respectively. We could then develop the sufficient conditions for the utility function $f(\alpha)$ to satisfy Theorem III.1 as developed for Table I. Given the specific values of $lambda_1$ and λ_2, one possible utility function will be in the form of $f(|S|) = -11|S|^2 + 20|S| + 41$.

B. *Cost function for Table II*

Table II
EXAMPLE 2 FOR ITEM PROBABILITY DISTRIBUTION

Hypotheses	e_1	e_2	e_3	e_4	e_5	e_6
H_1	0.10	0.13	0.23	0.28	0.17	0.19
H_2	0.19	0.12	0.20	0.20	0.08	0.21

Following the same method, we could easily show that the same utility function $f(|S|) = -11|S|^2 - +20|S| + 41$ satisfies the sufficient conditions gathered in Theorem III.1.

C. *Solution Comparison*

The results of running Algorithm 1 on the two scenarios have been gathered in Table III. Each cell represents the maximum overall utility achieved in either case by each method where once again it is obvious that the probability distribution plays a major role on the exactness of the solution compared to the utility function -which is the same in both cases. Also of interest is the negligible loss of utility at $\frac{48.76-48.7027}{48.7027} = 0.0012$ while a desirable cost reduction from NP complex to polynomial has taken place.

Table III
SOLUTION EXACTNESS COMPARISON FOR PROBLEM FORMULATED IN EQ.(9)

scenario	exhaustive search solution	submodular solution
1	47.68	47.68
2	48.7027	48.76

REFERENCES

[1] Y. Chen, C. Suh, and A. J. Goldsmith, "Information recovery from pairwise measurements," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5881–5905, Oct 2016.

[2] E. Zheleva and L. Getoor, *Preserving the Privacy of Sensitive Relationships in Graph Data*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 153–171.

[3] F. Bayat and S. Wei, "Partition of random items: Tradeoff between binning utility and meta information leakage," in *ICT*. Accpeted, 2018.

[4] F. Laforet, E. Buchmann, and K. B?hm, "Individual privacy constraints on time-series data," *Information Systems*, vol. 54, pp. 74 – 91, 2015.

[5] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the internet of things: threats and challenges," *Security and Communication Networks*, vol. 7, no. 12, pp. 2728–2742, 2014.

[6] R. Santiago and F. B. Shepherd, "Multi-agent and multivariate submodular optimization," *CoRR*, vol. abs/1612.05222, 2016.

[7] J. Liao, L. Sankar, V. Y. F. Tan, and F. du Pin Calmon, "Hypothesis testing in the high privacy limit," in *Allerton*. IEEE, 2016, pp. 649–656.

[8] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*. New York, NY, USA: John Wiley & Sons, Inc., 1998.

[9] M. Resende and C. Ribeiro, "Greedy randomized adaptive search procedures," in *Handbook of Metaheuristics*, F. Glover and G. Kochenberger, Eds. Kluwer Academic Publishers, 2003, pp. 219–249.

[10] D. Johnstone and D. Lindley, "Mean?variance and expected utility: The borch paradox," *Statist. Sci.*, vol. 28, no. 2, pp. 223–237, 05 2013.

VI. APPENDICES

A. *Theorem III.1*

Proof. By first imposing the diminishing returns property to $F(S_j)$ as formulated in Eq.(14) we see it necessary to have:

$$\begin{gathered}
\alpha_{x_1}\{-P(G_1)g(S_B^{(1)} + \{x\}^{(1)}, S_A^{(1)} + \{x\}^{(1)}) \\
+\lambda_1(\log(\alpha_{B_2} + \alpha_{x_2}) - \log(\alpha_{A_2} + \alpha_{x_2})) \\
+(P(G_1)\lambda_2 + \lambda_1)(\log(\alpha_{B_1} + \alpha_{x_1}) - \log(\alpha_{A_1} + \alpha_{x_1}))\} \\
+\alpha_{x_2}\{-P(G_2)g(S_B^{(2)} + \{x\}^{(2)}, S_A^{(2)} + \{x\}^{(2)}) \\
+\lambda_1(\log(\alpha_{B_1} + \alpha_{x_1}) - \log(\alpha_{A_1} + \alpha_{x_1})) \\
+(P(G_2)\lambda_2 + \lambda_1)(\log(\alpha_{B_2} + \alpha_{x_2}) - \log(\alpha_{A_2} + \alpha_{x_2}))\}
\end{gathered}$$

$$+\alpha_{BA_1}\{-P(G_1)g(S_B^{(1)}+\{x\}^{(1)},S_B^{(1)})$$
$$+\lambda_1(\log{(\alpha_{B_2}+\alpha_{x_2})}-\log{\alpha_{B_2}})$$
$$+(P(G_1)\lambda_2+\lambda_1)(\log{(\alpha_{B_1}+\alpha_{x_1})}-\log{\alpha_{B_1}})\}$$
$$+\alpha_{BA_2}\{-P(G_2)g(S_B^{(2)}+\{x\}^{(2)},S_B^{(2)})$$
$$+\lambda_1(\log{(\alpha_{B_1}+\alpha_{x_1})}-\log{\alpha_{B_1}})$$
$$+(P(G_2)\lambda_2+\lambda_1)(\log{(\alpha_{B_2}+\alpha_{x_2})}-\log{\alpha_{B_2}})\}$$
$$+\alpha_{A_1}\{-P(G_1)q(S_B^{(1)}+\{x\}^{(1)},S_B^{(1)},S_A^{(1)}+\{x\}^{(1)},S_A^{(1)})$$
$$+\lambda_1\log{(\frac{\alpha_{A_2}}{\alpha_{A_2}+\alpha_{x_2}}\frac{\alpha_{B_2}+\alpha_{x_2}}{\alpha_{B_2}})}$$
$$+(P(G_1)\lambda_2+\lambda_1)\log{(\frac{\alpha_{A_1}}{\alpha_{A_1}+\alpha_{x_1}}\frac{\alpha_{B_1}+\alpha_{x_1}}{\alpha_{B_1}})}\}$$
$$+\alpha_{A_2}\{-P(G_2)q(S_B^{(2)}+\{x\}^{(2)},S_B^{(2)},S_A^{(2)}+\{x\}^{(2)},S_A^{(2)})$$
$$+\lambda_1\log{(\frac{\alpha_{A_1}}{\alpha_{A_1}+\alpha_{x_1}}\frac{\alpha_{B_1}+\alpha_{x_1}}{\alpha_{B_1}})}$$
$$+(P(G_2)\lambda_2+\lambda_1)\log{(\frac{\alpha_{A_2}}{\alpha_{A_2}+\alpha_{x_2}}\frac{\alpha_{B_2}+\alpha_{x_2}}{\alpha_{B_2}})}\}\geq 0 \tag{20}$$

Now we attempt to find the sufficient conditions for Inequality (20) to hold true.

If we assume $g(S_j,S_w)\leq 0, S_w\subseteq S_j$ we will satisfy the positivity of α_{x_1}, α_{x_2}, α_{BA_1}, and α_{BA_2} coefficients.

It could further be seen that if we define

$$\omega^{-1}=min(\frac{\alpha_{A_1}}{\alpha_{A_1}+\alpha_{x_1}}\frac{\alpha_{B_1}+\alpha_{x_1}}{\alpha_{B_1}}) \tag{21}$$

and

$$\chi^{-1}=min(\frac{\alpha_{A_2}}{\alpha_{A_2}+\alpha_{x_2}}\frac{\alpha_{B_2}+\alpha_{x_2}}{\alpha_{B_2}}) \tag{22}$$

then as long as $q(S_j,S_w,S_r,S_t)\leq 0, S_w\subseteq S_j, S_t\subseteq S_r, S_r\subseteq S_j$ and

$$|q|\geq$$
$$max\{\lambda_2\log\omega+\frac{\lambda_1}{P(G_1)}\log{(\omega\chi)},$$
$$\lambda_2\log\chi+\frac{\lambda_1}{P(G_2)}\log{(\omega\chi)}\} \tag{23}$$

we will have satisfied the diminishing returns property inequality. In the next step we attempt to calculate the values of ω and χ. We further note that ω and χ represent the same set functions under different hypotheses, we could thus find the maximum value of either and then simply change the hypotheses index to derive the other. In the rest of this proof, we aim to find the maximal value of ω.

Since we are aiming to detect a lower bound over 3 interconnected variables α_{x_1}, α_{BA_1} and α_{A_1}, we need to show that the Hessian of the function ω is always negative meaning the value we find for ω is a maximum. However, it could be seen that such a relationship does not necessarily hold true. We thus choose to find an upper bound for such a lower bound rather than deal with it directly. We will have:

$$J(\alpha_{A_1},\alpha_{x_1})=1+\frac{\alpha_{x_1}\alpha_{BA_1}}{\alpha_{A_1}(\alpha_{A_1}+\alpha_{BA_1}+\alpha_{x_1})}$$
$$\leq 1+\frac{\alpha_{x_1}}{\alpha_{A_1}(\alpha_{A_1}+\alpha_{BA_1}+\alpha_{x_1})}\leq 1+\frac{\alpha_{x_1}}{\alpha_{A_1}(\alpha_{A_1}+\alpha_{x_1})}$$
$$=1+\frac{1}{\alpha_{A_1}}-\frac{1}{\alpha_{A_1}+\alpha_{x_1}}=J'(\alpha_{A_1},\alpha_{x_1}) \tag{24}$$

We now aim to maximize $J'(\alpha_{A_1},\alpha_{x_1})$. To do so, we first need to show that the Hessian of J' is always negative. We will have:

$$J'_{11}=\frac{\partial^2 J'}{\partial{\alpha_{A_1}}^2}=\frac{1}{{\alpha_{A_1}}^2}-\frac{2}{(\alpha_{A_1}+\alpha_{x_1})^3}$$
$$J'_{12}=J'_{21}\frac{\partial^2 J'}{\partial\alpha_{A_1}\partial\alpha_{x_1}}=-\frac{2}{(\alpha_{A_1}+\alpha_{x_1})^3}$$
$$J'_{22}=\frac{\partial^2 J'}{\partial{\alpha_{x_1}}^2}=-\frac{2}{(\alpha_{A_1}+\alpha_{x_1})^3} \tag{25}$$

It could then be seen that $J'_{11}J'_{22}-J'_{12}J'_{21}\leq 0$ meaning the function J' has a global maximum that could be found. To find this maximum we use the Lagrange Multiplier method where we hope to maximize:

$$\mathcal{L}=1+\frac{1}{\alpha_{A_1}}-\frac{1}{\alpha_{A_1}+\alpha_{x_1}}+$$
$$\lambda(\alpha_{A_1}+\alpha_{x_1}-K) \tag{26}$$

It turns out that J' is maximized when we have $\alpha_{x_1}=\pi_{1_1}$ and $\alpha_{A_1}=\pi_{M_1}$ and the resulting J' is:

$$J'(\pi_{M_1},\pi_{1_1})=1+\frac{\pi_{1_1}}{\pi_{M_1}(\pi_{1_1}+\pi_{M_1})} \tag{27}$$

Finally by reiterating ω using the new values we could see that

$$\omega\leq 1+\frac{\pi_{1_1}}{\pi_{M_1}(\pi_{1_1}+\pi_{M_1})} \tag{28}$$

It follows that

$$\chi\leq 1+\frac{\pi_{1_2}}{\pi_{M_2}(\pi_{1_2}+\pi_{M_2})} \tag{29}$$

Thus, for the function $F(S_j)$ as described in Eq.(14), we have the following sufficient conditions:

$$(1)\quad g(S_j,S_w)\leq 0$$
$$(2)\quad q(S_j,S_w,S_r,S_t)\leq 0$$
$$(3)\quad |q(S_j,S_w,S_r,S_t)|\geq$$
$$max\{\lambda_2\log\omega'+\frac{\lambda_1}{P(G_1)}\log{(\omega'\chi')},$$
$$\lambda_2\log\chi'+\frac{\lambda_1}{P(G_2)}\log{(\omega'\chi')}\}$$
$$,\omega'=1+\frac{\pi_{1_1}}{\pi_{M_1}(\pi_{1_1}+\pi_{M_1})},\chi'=1+\frac{\pi_{1_2}}{\pi_{M_2}(\pi_{1_2}+\pi_{M_2})}$$

Now we aim to carry out the same process for $F(S_j)$ as formulated in Eq.(18). By imposing the diminishing returns property, we see it necessary to have:

$$\alpha_{x_1}\{-P(G_1)g(S-S_A^{(1)}-\{x\}^{(1)},S-S_B^{(1)}-\{x\}^{(1)})$$
$$+\lambda_1(\log{(1-\alpha_{A_2}-\alpha_{x_2})}-\log{(1-\alpha_{B_2}-\alpha_{x_2})})$$
$$+(P(G_1)\lambda_2+\lambda_1)$$
$$(\log{(1-\alpha_{A_1}-\alpha_{x_1})}-\log{(1-\alpha_{B_1}-\alpha_{x_1})})\}$$
$$+\alpha_{x_2}\{-P(G_2)g(S-S_A^{(2)}-\{x\}^{(2)},S-S_B^{(2)}-\{x\}^{(2)})$$
$$+\lambda_1(\log{(1-\alpha_{A_1}-\alpha_{x_1})}-\log{(1-\alpha_{B_1}-\alpha_{x_1})})$$
$$+(P(G_2)\lambda_2+\lambda_1)$$
$$(\log{(1-\alpha_{A_2}-\alpha_{x_2})}-\log{(1-\alpha_{B_2}-\alpha_{x_2})})\}$$

$$\alpha_{BA_1}\{-P(G_1)g(S-S_B^{(1)}, S-S_B^{(1)}-\{x\}^{(1)})$$
$$+\lambda_1(\log(1-\alpha_{B_2})-\log(1-\alpha_{B_2}-\alpha_{x_2}))$$
$$+(P(G_1)\lambda_2+\lambda_1)$$
$$(\log(1-\alpha_{B_1})-\log(1-\alpha_{B_1}-\alpha_{x_1}))\}$$
$$+\alpha_{BA_2}\{-P(G_2)g(S-S_B^{(2)}, S-S_B^{(2)}-\{x\}^{(2)})$$
$$+\lambda_1(\log(1-\alpha_{B_1})-\log(1-\alpha_{B_1}-\alpha_{x_1}))$$
$$+(P(G_2)\lambda_2+\lambda_1)$$
$$(\log(1-\alpha_{B_2})-\log(1-\alpha_{B_2}-\alpha_{x_2}))\}$$

$$+(1-\alpha_{A_1})\{-P(G_1)q(S-S_A^{(1)}, S-S_A^{(1)}-\{x\}^{(1)},$$
$$S-S_B^{(1)}, S-S_B^{(1)}-\{x\}^{(1)})$$
$$+\lambda_1\log(\frac{1-\alpha_{A_2}}{1-\alpha_{A_2}-\alpha_{x_2}}\frac{1-\alpha_{B_2}-\alpha_{x_2}}{1-\alpha_{B_2}})$$
$$+(P(G_1)\lambda_2+\lambda_1)\log(\frac{1-\alpha_{A_1}}{1-\alpha_{A_1}-\alpha_{x_1}}\frac{1-\alpha_{B_1}-\alpha_{x_1}}{1-\alpha_{B_1}})\}$$
$$+(1-\alpha_{A_2})\{-P(G_2)q(S-S_A^{(2)}, S-S_A^{(2)}-\{x\}^{(2)},$$
$$S-S_B^{(2)}, S-S_B^{(2)}-\{x\}^{(2)})$$
$$+\lambda_1\log(\frac{1-\alpha_{A_1}}{1-\alpha_{A_1}-\alpha_{x_1}}\frac{1-\alpha_{B_1}-\alpha_{x_1}}{1-\alpha_{B_1}})$$
$$+(P(G_2)\lambda_2+\lambda_1)\log(\frac{1-\alpha_{A_2}}{1-\alpha_{A_2}-\alpha_{x_2}}\frac{1-\alpha_{B_2}-\alpha_{x_2}}{1-\alpha_{B_2}})\}$$
$$\geq 0 \quad (30)$$

Now we attempt to find the sufficient conditions for Inequality (30) to hold true.

Once again, if we assume $g(S_j, S_w) \leq 0, S_w \subseteq S_j$ we will satisfy the positivity of α_{x_1}, α_{x_2}, α_{BA_1}, and α_{BA_2} coefficients.

It could further be seen that if we define

$$\omega_2^{-1} = min(\frac{1-\alpha_{A_1}}{1-\alpha_{A_1}-\alpha_{x_1}}\frac{1-\alpha_{B_1}-\alpha_{x_1}}{1-\alpha_{B_1}}) \quad (31)$$

and

$$\chi_2^{-1} = min(\frac{1-\alpha_{A_2}}{1-\alpha_{A_2}-\alpha_{x_2}}\frac{1-\alpha_{B_2}-\alpha_{x_2}}{1-\alpha_{B_2}}) \quad (32)$$

then as long as $q(S_j, S_w, S_r, S_t) \leq 0, S_w \subseteq S_j, S_t \subseteq S_r, S_r \subseteq S_j$ and

$$|q| \geq$$
$$max\{\lambda_2\log\omega_2 + \frac{\lambda_1}{P(G_1)}\log(\omega_2\chi_2),$$
$$\lambda_2\log\chi_2 + \frac{\lambda_1}{P(G_2)}\log(\omega_2\chi_2)\} \quad (33)$$

we will have satisfied the diminishing returns property inequality. In the next step we attempt to calculate the values of ω_2 and χ_2. We further note that ω_2 and χ_2 represent the same set functions under different hypotheses, we could thus find the maximum value of either and then simply change the hypotheses index to derive the other. In the rest of this proof, we aim to find the maximal value of ω_2.

Once again, since we are aiming to detect a lower bound over 3 interconnected variables α_{x_1}, α_{BA_1} and α_{A_1}, we need to show that the Hessian of the function ω_2 is always negative meaning the value we find for ω_2 is a maximum. However, it could once more be seen that such a relationship does not necessarily hold true. We thus choose to find an upper bound for such a lower bound rather than deal with it directly again. We will have:

$$J_2(\alpha_{A_1}, \alpha_{x_1}) = 1 + \frac{\alpha_{x_1}\alpha_{BA_1}}{(1-\alpha_{A_1})(1-\alpha_{A_1}-\alpha_{BA_1}-\alpha_{x_1})}$$
$$\leq 1 + \frac{\alpha_{x_1}}{(1-\alpha_{B_1})(1-\alpha_{B_1}-\alpha_{x_1})}$$
$$= 1 + \frac{1}{1-\alpha_{B_1}-\alpha_{x_1}} - \frac{1}{1-\alpha_{B_1}} = J_2'(\alpha_{B_1}, \alpha_{x_1}) \quad (34)$$

Following the same logical steps, we could see that J_2' is maximized when we have $\alpha_{x_1} = \pi_{1_1}$ and $\alpha_{B_1} = 0$ and the resulting J_2' is:

$$J_2'(0, \pi_{1_1}) = 1 + \frac{\pi_{1_1}}{1-\pi_{1_1}} \quad (35)$$

Finally by reiterating ω_2 using the new values we could see that

$$\omega_2 \leq 1 + \frac{\pi_{1_1}}{1-\pi_{1_1}} \quad (36)$$

It follows that

$$\chi_2 \leq 1 + \frac{\pi_{1_2}}{1-\pi_{1_2}} \quad (37)$$

Thus, for the function $F(S_j)$ as described in Eq.(18), we have the following sufficient conditions:

$$(1) \quad g(S_j, S_w) \leq 0$$
$$(2) \quad q(S_j, S_w, S_r, S_t) \leq 0$$
$$(3) \quad |q(S_j, S_w, S_r, S_t)| \geq max$$
$$\{\lambda_2\log\omega_2' + \frac{\lambda_1}{P(G_1)}\log(\omega_2'\chi_2'),$$
$$\lambda_2\log\chi_2' + \frac{\lambda_1}{P(G_2)}\log(\omega_2'\chi_2')\}$$
$$,\omega_2' = 1 + \frac{\pi_{1_1}}{1-\pi_{1_1}}$$
$$,\chi_2' = 1 + \frac{\pi_{1_2}}{1-\pi_{1_2}} \quad (38)$$

Now that we have found the set of sufficient conditions for $F(S_j)$ in both cases of Eq.(14) and Eq.(18), we can find

the overall set of $F(S_j)$ in all cases as:

$$
\begin{aligned}
&(1) \quad g(S_j, S_w) \leq 0 \\
&(2) \quad q(S_j, S_w, S_r, S_t) \leq 0 \\
&(3) \quad |q(S_j, S_w, S_r, S_t)| \geq max(Z_1, Z_2, Z_3, Z_4), \\
&Z_1 = \lambda_2 \log \omega' + \frac{\lambda_1}{P(G_1)} \log(\omega' \chi'), \\
&Z_2 = \lambda_2 \log \chi' + \frac{\lambda_1}{P(G_2)} \log(\omega' \chi'), \\
&Z_3 = \lambda_2 \log \omega_2' + \frac{\lambda_1}{P(G_1)} \log(\omega_2' \chi_2'), \\
&Z_4 = \lambda_2 \log \chi_2' + \frac{\lambda_1}{P(G_2)} \log(\omega_2' \chi_2'), \\
&\omega' = 1 + \frac{\pi_{1_1}}{\pi_{M_1}(\pi_{1_1} + \pi_{M_1})}, \\
&\chi' = 1 + \frac{\pi_{1_2}}{\pi_{M_2}(\pi_{1_2} + \pi_{M_2})}, \\
&\omega_2' = 1 + \frac{\pi_{1_1}}{1 - \pi_{1_1}}, \\
&\chi_2' = 1 + \frac{\pi_{1_2}}{1 - \pi_{1_2}}
\end{aligned}
\tag{39}
$$

$\square$